DOI: 10.1111/fmii.12126

ORIGINAL ARTICLE

FINANCIAL MARKETS, INSTITUTIONS & INSTRUMENTS

WILEY

The time-varying performance of UK analyst recommendation revisions: Do market conditions matter?

Chen Su[1] | **Hanxiong Zhang**[2] | **Robert S. Hudson**[3]

[1]Newcastle University Business School

[2]Sheffield University Management School

[3]Hull University Business School

Correspondence
Chen Su, Newcastle University Business School, 5 Barrack Road, Newcastle upon Tyne NE1 4SE, UK.
Email: chen.su@ncl.ac.uk

Funding information
Newcastle University Business School; Faculty of Humanities and Social Sciences, Newcastle University, Grant/Award Number: OSR/0372/FR19/0001

Abstract

This study examines the time-varying performance of investment strategies following analyst recommendation revisions in the UK stock market, with specific emphasis on the impact of changing market conditions. We find a negative relationship between the recommendation performance and market conditions as measured in terms of past market return and market volatility. In particular, the *upgrade* (*downgrade*) portfolio generates significantly positive (negative) net abnormal returns in bad market conditions (e.g., the *dot-com* bubble burst in 2000 and the *credit* crisis in 2007), but not in other periods of time. Moreover, our non-temporal threshold regression analysis shows that the reported negative relationship disappears when market conditions become better, i.e., when the past market return (market volatility) is higher (lower) than a certain level, indicating the importance of taking non-linearity into account in the long sample period as examined in this study. Our time-series bootstrap simulations further confirm that the superior recommendation performance in bad market conditions is not due to random chance; analysts have certain skills in making valuable up/downward revisions in bad markets.

KEYWORDS

analyst recommendation revisions, bootstrap simulations, market conditions, non-temporal threshold regression model

JEL CLASSIFICATION

G11, G12, G14, G24

1 | INTRODUCTION

There are numerous studies on the information role of financial analysts, while the existing literature mostly ignores the question of whether the performance of their stock recommendations is related to the state of the economy (Chang & Choi, 2017; Loh & Stulz, 2018). Loh and Stulz (2018) provide empirical evidence showing that analysts' advice (e.g., stock recommendations and earnings forecasts) is more valuable in bad market conditions. They further argue that, in bad market conditions, analysts work harder and market investors rely more on analysts' advice than on other information sources. This argument, however, does not seem to align with the common sense view that bad market conditions, such as financial crises and recessions, usually give rise to increased uncertainty, making it much harder for analysts to make accurate stock recommendations (see, Amiram, Landsman, Owens, & Stubben, 2018; Bloom, 2009; Chopra, 1998). For example, Barber, Lehavy, McNichols, and Trueman (2003) report that in the years of 2000 and 2001 after the *dot-com* bubble burst, the most (least) favorable analyst recommendations lead to an average annualized abnormal return of –7.06% (13.44%).

In contributing to this debate, we shed fresh light on the time-varying performance of UK analyst recommendation revisions, with specific emphasis on the impact of market conditions. Our particular attention to the UK stock market is motivated by the following considerations. Specifically, despite the existence of extensive analyst research in the US market, empirical evidence on the performance of analyst recommendations remains mixed and far from conclusive (see relevant literature in Section 2). Jegadeesh and Kim (2006, p. 275) call for an in-depth investigation into other developed markets to provide "a more comprehensive picture of the extent to which the unique skills of analysts are useful for investors". The UK stock market, as a highly developed and sophisticated market, offers an appropriate setting to examine the recommendation performance and its implication on market efficiency. In particular, its institutional settings and trading practices are partially different from and independent of those in the US market;[1] as a result, the existing US evidence may not justify the UK investment practices. However, there is surprisingly little related analyst research in the UK stock market, except for two published studies that report conflicting results, using very small samples over relatively short time periods. For example, Dimson and Fraletti (1986) examine an unpublished sample of 1,649 telephone recommendations made by one UK broker in 1983 and find no significant abnormal returns for the recommended stocks. Ryan and Taffler (2006, p. 372) argue that the sample of analyst recommendations employed in Dimson and Fraletti (1986) is made by "a single UK brokerage house only and is biased towards large capitalization stocks". Ryan and Taffler (2006) investigate 2,506 analyst recommendation revisions made by six London-based brokers from December 1993 to June 1995, showing that stock prices are significantly affected by the changes in analyst recommendations.

This study examines a comprehensive sample of 70,220 UK analyst recommendation revisions over the period January 1995 to June 2013. As such, our sample is much larger than has been employed in the two prior UK studies; the long sample period also helps identify the evolution of the recommendation performance in different market conditions. Specifically, we construct an *upgrade* portfolio, including all upgrades to buy-related stock recommendations from previous sell/hold-related stock recommendations, as well as a *downgrade* portfolio, including all downgrades to sell/hold-related stock recommendations from previous buy-related stock recommendations. The *up/downgrade* portfolio is updated daily; for each up/downward revision, the recommended stock enters the *up/downgrade* portfolio at the close of trading on the day the revision is announced and then remains in the portfolio for up to five trading days (one week). Like Barber, Lehavy, McNichols, and Trueman (2001), we take an investor-oriented, calendar-time perspective to track the evolution of the portfolio performance on a rolling window basis.[2] This allows us to directly measure the time-varying performance of the *up/downgrade* portfolios using the intercepts derived from various asset pricing models and to estimate portfolio turnover. Consequently, we are able to determinate whether the *up/downgrade* portfolio can generate statistically significant net abnormal returns over time, after taking transaction costs into account (see, also, Barber et al., 2001). Accordingly, our empirical investigations proceed in three major parts using various alternative regression models.

In the first part of our investigations, our rolling window analysis results identify that the performance of the *upgrade* and *downgrade* portfolios varies considerably over time. In particular, the *upgrade* (*downgrade*) portfolio generates significantly positive (negative) net abnormal returns in two periods of bad market conditions, i.e., the *dot-com* bubble burst in 2000 and the *credit* crisis in 2007, but not in the rest of the sample period, in line with Loh and Stulz (2018). For example, the significantly positive daily net abnormal returns to the *upgrade* portfolio are observed over the period June 2000 to May 2002, ranging from 3.454 to 5.414 basis points (8.704% to 13.643% annualized; p-value ≤ 0.05) and over the period June 2007 to May 2008, ranging from 3.300 to 4.948 basis points (8.316% to 12.469% annualized; p-value ≤ 0.05). Similarly, the *downgrade* portfolio generates significantly negative daily net abnormal returns over the period June 2000 to March 2003, ranging from –3.970 to –7.255 basis points (–10.004% to –18.283% annualized; p-value ≤ 0.05) and over the period July 2007 to December 2009, ranging from –2.661 to –7.023 basis points (–6.706% to –17.698% annualized; p-value ≤ 0.05).

In the second part, we follow Derrien and Womack (2003) to measure market conditions as a continuous variable, in terms of past market return or market volatility (see, also, Chang & Choi, 2017). We run the ordinary least squares (OLS) regressions to formally test the impact of market conditions on the recommendation performance. Our OLS regression results show that the *up/downgrade* portfolio leads to superior performance in the face of more uncertainty shocks (i.e., lower market return and/or higher market volatility). For example, a decrease (increase) of 1% on the daily market return (market volatility), on average, gives rise to an increase of 2.780% (2.510%) in the performance of the *upgrade* portfolio. Similar evidence is found for the performance of the *downgrade* portfolio, in support of the findings in the first part of our investigations.

A recent study of Chan, Hansen, and Timmermann (2017) argues that it could be difficult for standard linear regression models to evaluate market properties during a long-term time period including various market conditions. In the third part of our empirical investigations, we employ the non-temporal threshold testing procedure originally proposed by Hansen (2000) to explore whether the dynamic recommendation performance is sensitive to changing market conditions. The non-temporal threshold model allows for regime-switching and identifying different market conditions endogenously. It also allows the non-linear effect to be driven by observable variables, but the number and value of thresholds are unknown *a priori* (see, Chan et al., 2017). Our non-temporal analysis confirms the negative relationship between the recommendation performance and market conditions, e.g., the superior recommendation performance in bad market conditions, consistent with those found in the first and second parts of our empirical investigations. However, once market conditions become better and exceed a certain level, the observed negative relationship disappears. Our results imply the importance of taking non-linearity into account when analyzing the impact of market conditions on the recommendation performance in the long sample period as examined in this study.

This study contributes to the analyst literature in several ways. First, the existing analyst research attributes the value of analyst recommendations to various factors, e.g., conflicts of interest, analyst reputation, timing, momentum, herding, and so on (see more details in Section 2), while the impact of changing market conditions has received comparatively little attention (see, e.g., Chang & Choi, 2017; Loh & Stulz, 2018). To the best of our knowledge, this is the first study to examine the relationship between market conditions and the recommendation performance in the UK stock market using various alternative methodological approaches, in particular, the non-temporal threshold regression model that has not been employed in prior analyst research. In addition, we provide some out-of-sample evidence to complement the existing analyst literature–the superior recommendation performance in bad market conditions in the UK market–that is not only statistically significant, but also economically meaningful to market investors. From an investor's perspective, it is possible for market investors to make profits by purchasing (short selling) stocks with upward (downward) revisions in bad market conditions with high uncertainty, even after accounting for transaction costs. Finally, our results hold up against an array of robustness checks, including bootstrap simulations and various asset pricing models. For example, to rule out the concern that the observed superior recommendation performance in bad market conditions might be spurious (see, Barber et al., 2001; Fama, 1998), we develop a time-series bootstrap simulation method to distinguish analysts' skill from luck. Our bootstrap simulations confirm that the observed superior recommendation performance in bad market conditions is not due to analysts' luck (i.e., random chance), but due to analysts' skill. That is, financial analysts possess sufficient skill to make valuable up/downward revisions in bad market conditions. Also, our results are robust to various single- and multi-factor assets pricing models, ruling out Barber et al. (2001) concern that the observed superior recommendation performance in bad market conditions could be due to a poor model of asset pricing.

The remainder of this paper is organized as follows. The next section reviews related analyst literature and develops our major hypotheses. Section 3 describes our data and methodology. Section 4 presents empirical results, followed by various robustness checks in Section 5. The final section concludes.

2 | RELATED ANALYST LITERATURE AND HYPOTHESES DEVELOPMENT

Financial analysts play an important role as information intermediaries in the capital market, collecting and analyzing a wide variety of market, industry, and firm-specific information, and then making stock recommendations to investors. A significant body of academic research has been devoted to investigating the question of whether investors can profit from analyst recommendations since the seminal work of Cowles (1933), while there is a clear gap between theory and practice regarding this issue. For example, almost all brokerage houses make great efforts and spend large sums of money on security analysis; as a result, their analyst recommendations have been widely used by market participants when making investment decisions. However, the semi-strong form of market efficiency implies that it is impossible for investors to make profits by using publicly available information. Grossman and Stiglitz (1980) argue that if prices fully reflect all publicly available information, then the use of analyst recommendations cannot generate superior returns, and brokerage houses should not spend massive money on security analysis, nor should investors have any incentives to pay for such costly information.

Much of the early evidence shows that investors are not able to add value to the market when they follow analyst recommendations. Dimson and Marsh (1984, p. 1259) summarize 27 early studies on the value of analyst recommendations documenting that "the profitability opportunities disclosed by these studies are, however, limited". In contrast, Stickel (1995) and Womack (1996) report that upgrades and downgrades are respectively accompanied by significantly positive and negative returns at the time of their announcements (see, also, Barber et al., 2001; Boni & Womack, 2006; Green, 2006; Jegadeesh, Kim, Krische, & Lee, 2004; Mokoaleli-Mokoteli, Taffler, & Agarwal, 2009; among others),[3] which seems not to be in accord with semi-strong form market efficiency. However, Barber et al. (2001) argue that these investment strategies require a great deal of trading and generate considerable transaction costs, suggesting that the apparent market inefficiency might not be easily exploitable by investors (see, also, Hall & Tacon, 2010; Jegadeesh et al., 2004; Mikhail, Walther, & Willis, 2004). Therefore, in this study, we evaluate the performance of the *up/downgrade* portfolio after accounting for transaction costs, that is, the portfolio performance is measured as the gross returns (estimated from various single- and multi-factor asset pricing models) less the estimated transaction costs multiplied by the corresponding daily portfolio turnover (see details in Subsection 3.3.2).

The existing research attributes the value of analyst recommendations to a variety of factors, such as conflicts of interest (see, Mehran & Stulz, 2007; Shen & Chih, 2009), analyst reputation (see, Emery & Li, 2009; Fang & Yasuda, 2009, 2014; Kucheev, Ruiz, & Sorensson, 2017), industry (see, Boni & Womack, 2006; Bradley, Gokkaya, & Liu, 2017; Merkley, Michaely, & Pacelli, 2017), timing (see, Green, 2006; Irvine, Lipson, & Puckett, 2007; Ivkovic & Jegadeesh, 2004), herding (see, Jegadeesh & Kim, 2010; Trueman, 1994), and so on. A very recent study by Loh and Stulz (2018) argues that the usefulness and performance of analysts could be dependent on bad market conditions, though little attention has been paid to this hypothesis. On the one hand, it is well known that bad market conditions tend to give rise to increased uncertainty (Bloom, 2009), so it should be more difficult for analysts to make accurate stock recommendations and earnings forecasts (see, e.g., Amiram et al., 2018; Chopra, 1998; Forbes, Murphy, O'Keeffe, & Su, 2020). Also, the decline in trading volume and hence broker profits in bad market conditions may reduce analysts' performance rewards, leading to a decrease in analysts' motivation. On the other hand, a theoretical model proposed by Kacperczyk and Seru (2007) suggests that investors rely more on analysts' advice than on other information sources. In particular, in bad market conditions with high uncertainty, investors' private information or information-processing ability becomes much noisier, making it more difficult for them to understand the information contained in financial documents and to assess the consequences of uncertainty shocks. High uncertainty, therefore, drives investor demand for information from financial analysts, as they are more experienced and informed and better able to evaluate uncertainty shocks in bad market conditions (see, Amiram et al., 2018; Kacperczyk, Van Nieuwerburgh, & Veldkamp, 2016; Loh & Stulz, 2011). To the extent that the role of analysts is to make sense of firms in the face of uncertainty, they should work harder in bad market condition due to career concerns (Loh & Stulz, 2018). For example, Merkley et al. (2017) show that analysts revise their earnings forecasts more frequently and release longer reports in bad market conditions.

If financial analysts can truly add value in bad market conditions, we may feel more confident in their overall role in predicting the nation's corporate prospects. We expect analysts to produce high quality stock recommendations to help investors understand the potential impact of uncertainty shocks in bad market conditions. Accordingly, we test the following main hypotheses relating to the proposition that analyst recommendation revisions are more valuable in bad market conditions. That is,

Hypothesis 1a: The average net abnormal return to the *upgrade* portfolio is significantly *positive* in bad market conditions, but is statistically insignificant in the rest of the sample period;

Hypothesis 1b: The average net abnormal return to the *downgrade* portfolio is significantly *negative* in bad market conditions, but is statistically insignificant in the rest of the sample period.

3 | DATA AND RESEARCH DESIGN

3.1 | Data and sample selection

We obtain real-time UK analyst recommendations from *the Morningstar Extracted Data File: Historic Broker Recommendations for UK Registered and UK Listed Companies*, uniquely created by *Morningstar Company Intelligence*.[4] Each stock recommendation record contains information on the name of the recommended firm, the name of the brokerage house issuing the recommendation, the starting and expiration dates of the recommendation, and a rating between 1 and 9 (1 = strong buy; 2 = buy; 3 = weak buy; 4 = weak buy/hold; 5 = hold; 6 = hold/sell; 7 = weak sell; 8 = sell; and 9 = strong sell). According to Altinkilic and Hansen (2009) and Altinkilic et al. (2013), we also exclude those stock recommendations made in the three days around quarterly earnings announcements (see, also, Loh & Stulz, 2011). We further require that the gap between the starting and expiration dates of the recommendation is less than one year to ensure that the brokerage house actively follows the recommended stock. The relevant financial data for the recommended firms are obtained from the London Share Price Database (LSPD).

Our initial sample is comprised of 384,165 publicly available analyst recommendations made by 144 brokerage houses on 2,905 distinct firms listed either on the London Stock Exchange (LSE) over the period January 1995 to June 2013. To allow for an easy and intuitive comparison with prior analyst research, we reclassify all original analyst recommendations into five categories: Strong Buys (1 & 2; 44.99%), Buys (3 & 4; 10.66%), Holds (5; 32.43%), Sells (6 & 7; 3.68%), and Strong Sells (8 & 9; 8.24%). The distribution of UK analyst recommendations in the *Morningstar* database is similar to that reported in prior US studies (see details in Appendix A).

Financial analysts, however, often leave their stock recommendations unchanged for long periods, with an average gap of 68 days between the starting and expiration dates in the UK (see Appendix A). Analyst recommendations thus become stale and less informed over time, potentially resulting in poor portfolio performance (see, Boni & Womack, 2006; Jegadeesh & Kim, 2006; Jegadeesh et al., 2004). Therefore, our study exclusively focuses on analyst recommendation revisions, which tend to convey more valuable information and have stronger predictive power.

TABLE 1 The transition matrix of UK analyst recommendation revisions

From old rating	To new rating: Strong Buy (1 & 2)	Buy (3 & 4)	Hold (5)	Sell (6 & 7)	Strong Sell (8 & 9)	Total	%
Strong Buy (1 & 2)	—	5,923	14,506	396	977	21,802	31.05
Buy (3 & 4)	5,756	—	5,242	682	204	11,884	16.92
Hold (5)	12,745	5,267	—	2,596	4,927	25,535	36.36
Sell (6 & 7)	298	630	2,410	—	844	4,182	5.96
Strong Sell (8 & 9)	833	172	5,043	769	—	6,817	9.71
Overall	**19,632**	**11,992**	**27,201**	**4,443**	**6,952**	**70,220**	—
%	27.96	17.08	38.73	6.33	9.90	—	100.00

Note: This table presents the transition matrix of 70,220 UK analyst recommendation revisions over the period January 1995 to June 2013. All real-time analyst recommendations are obtained from *Morningstar Company Intelligence*. A rating of 1 reflects a strong buy, 2 a buy, 3 a weak buy, 4 a weak buy/hold, 5 a hold, 6 a hold/sell, 7 a weak sell, 8 a sell, and 9 a strong sell, which are reclassified into five categories: Strong Buy (1 & 2), Buy (3 & 4), Hold (5), Sell (6 & 7), and Strong Sell (8 & 9).

Table 1 presents the transition matrix of our final sample of 70,220 UK analyst recommendation revisions, i.e., 45.04% are Strong Buys and Buys, 38.73% are Holds, and 16.23% are Sells and Strong Sells. As such, our dataset is much larger than has been employed in two prior UK studies of Dimson and Fraletti (1986) and Ryan and Taffler (2006), while the long sample period of January 1995 to June 2013 allows us to observe the evolution of the portfolio performance in different market conditions. The inclusion of 23,235 (33.09%) analyst recommendation revisions on 1,244 dead firms in our sample also helps alleviate the potential survivorship bias.

3.2 | Portfolio construction

To evaluate the performance of calendar-time investment strategies based on analyst recommendation revisions in the UK market, we construct two portfolios: (i) an *upgrade* portfolio, consisting of all stocks with upward revisions to Strong Buy or Buy recommendations from previous Strong Sell, Sell, or Hold recommendations; and (ii) a *downgrade* portfolio, consisting of all stocks with downward revisions to Strong Sell, Sell, or Hold recommendations from previous Strong Buy or Buy recommendations. Specifically, the *upgrade* portfolio is updated daily; for each upward revision, the recommended stock enters the *upgrade* portfolio at the close of trading on the day the revision is announced. If an upward revision is announced on a non-trading day, the recommended stock is added into the *upgrade* portfolio at the close of the next trading day. The recommended stock then remains in the portfolio for up to five trading days (one week) after the revision is announced to explore the short-term market reaction to analyst recommendation revisions.[5],[6] The *downgrade* portfolio is constructed in an analogous fashion.

Panel A of Table 2 presents the distribution of upward and downward revisions in the *upgrade* and *downgrade* portfolios, respectively, in each recommendation year. The total number of analyst recommendation revisions included in the *upgrade* and *downgrade* portfolios, 56,075 (= 25,701 + 30,374), appears to be less than 70,220, as shown in Table 1, which is not surprising, however. The upward revisions from Strong Sells to Sells, from Strong Sells to Holds, and from Sells to Holds are not included in the *upgrade* portfolio, as they can also be interpreted as negative recommendations. Similarly, the downward revisions from Strong Buys to Buys are not included in the *downgrade* portfolio, as they can also be interpreted as positive recommendations (see, also, Stickel, 1995). Panel B of Table 2 presents the distribution of upward and downward revisions in the *upgrade* and *downgrade* portfolios, respectively, in each industry category, according to the two-digit Industry Classification Benchmark (ICB) codes.[7]

TABLE 2 The distribution of UK analyst recommendation revisions in the *upgrade* and *downgrade* portfolios

	The *upgrade* portfolio				The *downgrade* portfolio			
	No. of firms covered	No. of brokerage houses	Average rating	No. of upward revisions	No. of covered firms	No. of brokerage houses	Average rating	No. of downward revisions
Panel A: The recommendation year								
1995	340	22	1.15	621	422	21	3.68	982
1996	603	37	1.20	1,498	585	34	3.54	1,715
1997	654	46	1.23	1,978	654	45	3.54	2,389
1998	660	37	1.28	1,936	654	39	3.59	2,395
1999	625	40	1.28	1,993	618	41	3.53	1,801
2000	513	39	1.27	1,419	512	39	3.45	1,496
2001	530	41	1.26	1,331	616	40	3.64	2,074
2002	567	41	1.24	1,341	569	42	3.74	1,480
2003	507	41	1.22	1,165	566	39	3.64	1,591
2004	578	43	1.24	1,499	558	47	3.65	1,818
2005	602	41	1.28	1,628	629	45	3.62	2,182
2006	575	44	1.28	1,608	600	44	3.56	1,834
2007	573	38	1.24	1,581	544	41	3.52	1,448
2008	488	41	1.21	1,204	548	39	3.67	1,680
2009	569	47	1.19	1,849	560	48	3.61	1,953
2010	470	39	1.20	1,145	436	36	3.43	1,192
2011	419	35	1.18	977	414	33	3.43	1,022
2012	340	28	1.16	665	409	29	3.49	957
2013 (January to June)	192	24	1.21	263	234	22	3.53	365
Panel B: The ICB industry category								
05 Oil & Gas	84	69	1.24	1,050	89	68	3.60	1,089
13 Chemicals	39	52	1.24	701	40	54	3.44	763
17 Basic Resources	74	55	1.22	621	72	58	3.67	767
23 Construction & Materials	66	54	1.26	929	64	58	3.61	1,022
27 Industrial Goods & Services	438	79	1.25	5,808	469	89	3.49	6,365
33 Automobiles & Parts	14	35	1.26	184	18	43	3.56	245
35 Food & Beverage	60	53	1.25	1,081	67	58	3.60	1,396
37 Personal & Household Goods	115	61	1.25	1,451	129	67	3.55	1,665
45 Health Care	103	67	1.22	948	114	66	3.52	1,161
53 Retail	130	71	1.25	2,782	143	75	3.75	4,005
55 Media	130	66	1.22	1,252	139	66	3.46	1,375
57 Travel & Leisure	143	70	1.19	2,462	149	76	3.55	2,944
65 Telecommunications	32	58	1.15	318	33	57	3.63	363
75 Utilities	32	44	1.23	551	29	45	3.64	666
83 Banks	14	49	1.23	606	15	48	3.84	806
85 Insurance	60	52	1.31	967	62	52	3.54	999

(Continues)

TABLE 2 (Continued)

Panel B: The ICB industry category								
87 Financial Services	213	70	1.27	2,173	215	71	3.52	2,376
95 Technology	211	74	1.17	1,817	233	75	3.62	2,367
Overall	**1,958**	**111**	**1.24**	**25,701**	**2,081**	**109**	**3.58**	**30,374**

Note: Panel A of this table presents the distribution of 25,701 upward (30,374 downward) changes in analyst recommendations in the *upgrade* (*downgrade*) portfolio over the period January 1993 to June 2013 by the recommendation year, in terms of the number of recommended firms, the number of brokerage houses, as well as the average rating and number of analyst recommendation revisions. Panel B of this table presents the distribution of all upward and downward revisions by the industry category, according to the two-digit ICB codes (see full details and description of the ICB at: http://www.icbenchmark.com). All real-time analyst recommendations are obtained from *Morningstar Company Intelligence*. A rating of 1 reflects a strong buy, 2 a buy, 3 a weak buy, 4 a weak buy/hold, 5 a hold, 6 a hold/sell, 7 a weak sell, 8 a sell, and 9 a strong sell, which are reclassified into five categories: Strong Buy (1 & 2), Buy (3 & 4), Hold (5), Sell (6 & 7), and Strong Sell (8 & 9). An *upgrade* portfolio consists of all upward revisions to Strong buy or Buy from previous Strong Sell, Sell, or Hold, while a *downgrade* portfolio consists of all downgrades to Strong Sell, Sell, or Hold from previous Strong Buy or Buy. The *upgrade* portfolio does not include upward revisions from Strong Sell to Hold, from Strong Sell to Sell, and from Sell to Hold, which can also be interpreted as negative recommendations, while the *downgrade* portfolio does not include downward revisions from Strong Buy to Buy, which can also be interpreted as positive recommendations. We report the average rating for analyst recommendation revisions based on the five-point rating scale.

3.3 | Portfolio performance evaluation

3.3.1 | The gross returns to the *up/downgrade* portfolio

Like Barber, Lehavy, and Trueman (2007), we apply an equal monetary investment in each analyst recommendation revision, and calculate the daily value-weighted return to the *upgrade* or *downgrade* portfolio on date t:

$$R_{p,t} = \left(\sum_{i=1}^{n_t} x_{i,t} \times R_{i,t} \right) \Big/ \sum_{i=1}^{n_t} x_{i,t}, \tag{1}$$

where $R_{i,t}$ represents the daily return for the recommended stock i on date t;[8] n_t represents the number of upward (downward) revisions in the *upgrade* (*downgrade*) portfolio p at the close of trading of the revision date through date $t-1$; $x_{i,t}$ represents the compounded daily return for the recommended stock i from the closing of trading on the revision date through date $t-1$ ($x_{i,t} = 1$ for a stock recommended on date $t-1$).[9]

The gross returns to the *upgrade* and *downgrade* portfolios are estimated using the intercept term of α_p derived from the capital asset pricing model (CAPM):

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_p \left(R_{m,t} - R_{f,t} \right) + \varepsilon_{p,t}, \tag{2}$$

where $R_{p,t}$ and $R_{m,t}$ represent the daily return on the portfolio and on the FTSE All-Share Index, respectively; $R_{f,t}$ represents the daily three-month UK T-bill rate; $\varepsilon_{p,t}$ represents the error term.

We estimate Eq. (2) repeatedly on a rolling window basis–a one-year window length rolling one trading day forward–to track the time-varying performance of the underlying variables over our long sample period. Specifically, the first rolling window is from January 2, 1995 to December 29, 1995, covering 252 trading days–the typical number of trading days in a year in the UK stock market. Then, a new observation (trading day) is added to the rolling window, while the first one observation is dropped, that is, we update the rolling window to include observations from January 3, 1995 to January 2, 1996, and so forth. In each rolling window, a significantly positive (negative) α_p indicates that the *upgrade* (*downgrade*) portfolio is profitable after controlling for market risk. This calculation, therefore, generates a time series of 4,420 daily gross returns to the *upgrade* or *downgrade* portfolio, over the period January 1996 to June 2013.

3.3.2 | The abnormal returns net of transaction costs

In this study, we focus on the average daily net abnormal returns to the *upgrade* and *downgrade* portfolios after taking transaction costs into account, that is, we consider markets to be inefficient only if the gross returns estimated from various asset pricing models are large enough to cover the size of transaction costs (see, Timmermann & Granger, 2004). Keim and Madhavan (1998) categorize transaction costs into explicit costs (e.g., brokerage commissions and taxes) and implicit costs (e.g., bid-ask spread and market impact of trading). Hudson, Dempsey, and Keasey (1996) show that the total round-trip transaction costs in the UK stock market for the most favored of investors is upward of 1.0%, including government stamp duty of 0.5%, negotiated brokerage commission of 0.1% (soft commissions could be zero if alternative services are offered in lieu of cash), and bid-ask spread of 0.5%. Based on a relatively cautious estimate of the average round-trip transaction costs in the UK for purchasing stocks at 1.5% and for short selling stocks at 3.0%,[10] we measure the round-trip transaction costs multiplied by the corresponding average daily portfolio turnover in each rolling window.

Specifically, the daily turnover for the portfolio on the trading date t is defined as the percentage of stocks in the portfolio as of the close of trading on date $t-1$ that have been sold off by the close of trading on date t. That is, like Barber et al. (2001), we measure the daily turnover as the percentage of the portfolio that has been moved into some other set of stocks on date t. For each stock i in portfolio p as of the close of trading on date $t-1$, we calculate its fraction of the portfolio, $G_{i,t}$, at the end of trading on date t without accounting for portfolio rebalancing:

$$G_{i,t} = \omega_{i,t-1} \times \left(1 + R_{i,t}\right) \Big/ \sum_{i=1}^{n_{p,t-1}} \omega_{i,t-1} \times \left(1 + R_{i,t}\right). \tag{3}$$

Then, $G_{i,t}$ is compared to the actual fraction $F_{i,t}$ that stock i makes up of portfolio p as of the close of trading on date t, after accounting for any portfolio rebalancing. Finally, the change in the percentage holding of each stock on date $t-1$ is summed, generating the portfolio turnover on date t:

$$TURNOVER_{p,t} = \sum_{i=1}^{n_{p,t}} \left|G_{i,t} - F_{i,t}\right|. \tag{4}$$

In each rolling window, we calculate the net abnormal returns as the gross returns less the estimated transaction costs multiplied by the corresponding daily portfolio turnover.

4 | EMPIRICAL RESULTS

In Subsection 4.1, our rolling window analysis results show significantly positive (negative) net abnormal returns to the *upgrade* (*downgrade*) portfolio in two periods of bad market conditions. In Subsection 4.2, we formally test the impact of changing market conditions on the performance of analyst recommendation revisions, confirming the existence of a significantly negative relationship between the recommendation performance and market conditions, measured as in terms of past market return and/or market volatility. In Subsection 4.3, our non-temporal threshold regression results confirm that the observed negative relationship disappears when market conditions become better, i.e., when past market return (market volatility) is higher (lower) than a certain level.

4.1 | Rolling window analysis results

Figure 1a (1b) illustrates the time-varying average daily abnormal returns net of transaction costs to the *upgrade* (*downgrade*) portfolio under the CAPM, along with the corresponding t-statistics. Specifically, the net abnormal returns to the *upgrade* and *downgrade* portfolios vary over time and, in particular, the *upgrade* (*downgrade*) portfolio generates significantly positive (negative) net abnormal returns in two periods of bad market conditions, i.e., the *dot-com* bubble burst

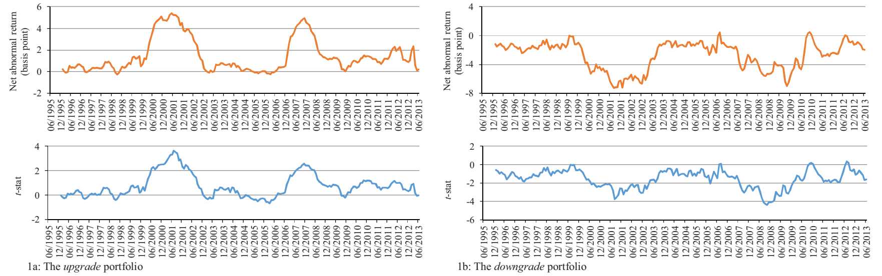

FIGURE 1 The time-varying daily net abnormal returns to the *up/downgrade* portfolio and the corresponding *t*-statistics under the CAPM, over the period January 1996 to June 2013 [Color figure can be viewed at wileyonlinelibrary.com]

in 2000 and the *credit* crisis in 2007. Figure 1a illustrates that the net abnormal returns to the *upgrade* portfolio initially fluctuate around zero, and then become significantly positive from June 2000 to May 2002, ranging from 3.454 to 5.414 basis points (8.704% to 13.643% annualized; *p*-value $\leq$ 0.05).[11] After that, the *upgrade* portfolio generates insignificant net abnormal returns up to June 2007, followed by significantly positive net abnormal returns from June 2007 to May 2008, ranging from 3.300 to 4.948 basis points (8.316% to 12.469% annualized; *p*-value $\leq$ 0.05). Also, Figure 1b illustrates that the *downgrade* portfolio generates significantly negative net abnormal returns in the two bad periods: (i) from June 2000 to March 2003, ranging from −3.970 to −7.255 basis points (−10.004% to −18.283% annualized; *p*-value $\leq$ 0.05); and (ii) from July 2007 to December 2009, ranging from −2.661 to −7.023 basis points (−6.706% to −17.698% annualized; *p*-value $\leq$ 0.05). The statistically significant net abnormal returns to the *downgrade* portfolio are not found in the remaining sample period.

Two important implications can be drawn from the patterns of Figures 1a & 1b. First, our results support Hypothesis 1a (1b) that the performance of the *upgrade* (*downgrade*) portfolio is significantly *positive* (*negative*) in bad market conditions, such as the *dot-com* bubble burst in 2000 and the *credit* crisis in 2007, but remains statistically insignificant in the remaining periods. We discuss the impact of market conditions on the recommendation performance in the next subsection in more detail. Second, from an investor's perspective, the significantly positive (negative) net abnormal returns to the *upgrade* (*downgrade*) portfolio are exploitable, that is, it is possible for investors to make profits by purchasing (short selling) stocks with publicly available upward (downward) revisions in bad market conditions, even after taking transaction costs into account.[12] This is confirmed by our bootstrap simulations in Section 5.

4.2 | The impact of market conditions on the recommendation performance

As the impact of market conditions is extremely time-period dependent and varies greatly relying on the length of the period chosen (Derrien & Womack, 2003), we examine various time periods that encompass the time frame before the portfolio is rebalanced. Thus, the market return variables are constructed for the one-month (*MktRet_1m*), two-month (*MktRet_2m*), and three-month (*MktRet_3m*) periods before the portfolio is rebalanced as an estimate of the buy-and-hold return on the FTSE All-Share Index. Also, a one-month *weighted* market return variable (*MktRet_w*) is constructed as a time weighted average buy-and-hold return of the corresponding market index return in the three months before the portfolio is rebalanced. The weights are 3 for the most recent month, 2 for the previous month, and 1 for the third month before the portfolio is rebalanced, based on the assumption that investors' perceptions take the last three months into account, but give heavier weight to more recent periods (see, Derrien & Womack, 2003). In addition, we choose four commonly used macroeconomic factors, including (i) the inflation rate computed from the consumer price index (*CPI*), (ii) the industrial production growth (*PROD*), (iii) the dividend yield (*DIV*), and (iv) the detrended short-term interest rate (*IR*), to control for the potential influence of macroeconomic situations and business cycles (see, e.g., Flannery & Protopapadakis, 2002; Hjalmarsson, 2010). Thus, we regress the daily net abnormal returns to the *upgrade* and *downgrade* portfolios, separately, on a set of macroeconomic variables and each market condition variable, along with two indicator variables of *YEAR* and *INDUSTRY* to control for the potential year and industry fixed effects, respectively:

$$\begin{aligned} Net\ Abnormal\ Returns &= \alpha + \beta_1 CPI + \beta_2 PROD + \beta_3 DIV + \beta_4 IR + \beta_5 \left[Market\ Return\ Variable\right] \\ &+ YEAR + INDUSTRY + \varepsilon. \end{aligned} \tag{5}$$

The left side of Table 3 shows that, in Regressions (2–5), all market return variables are significantly negative with the use of the net abnormal returns to the *upgrade* portfolio as the dependent variable under the CAPM. For example, in Regression (5), the market return variable of *MktRet_w* is statistically significant at the 5% level (t-stat = 2.24), while the coefficient of –2.165 also suggests that a decrease of 1% on the daily market return, on average, gives rise to an increase of 2.165% in the net abnormal return to the *upgrade* portfolio. In addition, the explanatory power of Regression (5), an adjusted R^2 of 0.240, is largely driven by the market return variable, as the adjusted R^2 is substantially reduced to 0.036 in Regression (1) when the market return variable is excluded.

Similar to Amiram et al. (2018), we further test the impact of market conditions in terms of market volatility (*MktVol_1m*), measured as the standard deviation of the daily return of the FTSE All-Share Index in the month before the portfolio is rebalanced, on the net abnormal returns to the *up/downgrade* portfolio. In addition to a set of macroeconomic variables, the *weighted* market return variable, and two indicator variables, included in Eq. (6), we introduce the market volatility variable:

$$\begin{aligned} Net\ Abnormal\ Returns &= \alpha + \beta_1 CPI + \beta_2 PROD + \beta_3 DIV + \beta_4 IR + \beta_5 MktRet\_w + \beta_6 MktVol\_1m \\ &+ YEAR + INDUSTRY + \varepsilon. \end{aligned} \tag{6}$$

We find that, similarly to the market return variable, the market volatility variable plays an important role in explaining the performance of the *upgrade* portfolio. For example, Panel A of Table 3 shows that, in Regression (6), the coefficient on *MktVol_1m*, 2.510 (t-stat = 2.78), is significantly positive at the 1% level, using the abnormal net returns to the *upgrade* portfolio as the dependent variable. The magnitude of the coefficient on *MktVol_1m* suggests that an increase of 1% in market volatility daily, on average, leads to an additional increase of around 2.510% in the net abnormal return to the *upgrade* portfolio. The inclusion of the market volatility variable improves the explanatory power of the model, i.e., the adjusted R^2 substantially increases to 0.290 in Regression (6).

Similarly, the right side of Table 3 shows that past market return and volatility also have statistically and economically significant impacts on the net abnormal returns to the *downgrade* portfolio.[13] In sum, the *up/downgrade* portfolio gives rise to superior performance in the face of lower market return and/or higher market volatility, making it is possible for investors to make profits by purchasing (short selling) stocks with publicly available upward (downward) revisions in bad market conditions.

4.3 | Non-temporal threshold regression results

Chan et al. (2017) argue that it could be difficult to model market properties over a long-term time period including various market conditions with the use of standard linear regression models. In the final part of our empirical analyses, we employ the non-temporal threshold testing procedure, originally proposed by Hansen (2000), to test whether the dynamic recommendation performance is sensitive to different market conditions, in terms of past market return or market volatility. The Hansen (2000) non-temporal threshold models are as follows:[14]

$$\begin{aligned} \left[Net\ Abnormal\ Returns\right]_t &= \alpha_B + \beta_B \left(R_{m,t} - R_{f,t}\right) + \varepsilon_{B,t} \quad if\ MktRet\_w_t \le x; \\ \left[Net\ Abnormal\ Returns\right]_t &= \alpha_G + \beta_G \left(R_{m,t} - R_{f,t}\right) + \varepsilon_{G,t} \quad if\ MktRet\_w_t > x, \end{aligned} \tag{7}$$

TABLE 3 The impacts of market conditions, in terms of past market return and volatility, on the performance of the *upgrade* and *downgrade* portfolios under the CAPM

	The *upgrade* portfolio						The *downgrade* portfolio					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
INTERCEPT	0.058	0.039	0.037	0.041	0.038	0.020	−0.098	−0.071	−0.069	−0.074	−0.064	−0.044
	(0.97)	(0.65)	(0.63)	(0.68)	(0.66)	(0.32)	(−0.93)	(−0.84)	(−0.78)	(−0.88)	(−0.79)	(−0.50)
CPI	−0.324[a]	−0.305[a]	−0.317[a]	−0.320[a]	−0.314[a]	−0.298[a]	−0.257[b]	−0.315[b]	−0.335[b]	−0.353[a]	−0.365[a]	−0.311[b]
	(−3.31)	(−3.07)	(−3.19)	(−3.23)	(−3.16)	(−3.10)	(−2.26)	(−2.26)	(−2.50)	(−2.74)	(−2.78)	(−2.29)
PROD	−0.129[c]	−0.136[c]	−0.131[c]	−0.130[c]	−0.132[c]	−0.124[c]	−0.164[b]	−0.130[b]	−0.139[a]	−0.132[a]	−0.126[b]	−0.128[b]
	(−1.76)	(−1.79)	(−1.74)	(−1.73)	(−1.75)	(−1.80)	(−2.00)	(−2.45)	(−2.74)	(−2.69)	(−2.53)	(−2.08)
DIV	0.588[a]	0.620[a]	0.588[a]	0.584[a]	0.589[a]	0.519[a]	0.257[b]	0.345[b]	0.316[c]	0.292[b]	0.326[b]	0.316[b]
	(3.31)	(3.44)	(3.30)	(3.27)	(3.31)	(3.16)	(1.99)	(2.13)	(1.92)	(1.97)	(2.02)	(2.26)
IR	−0.106[b]	−0.098[b]	−0.102[b]	−0.106[b]	−0.101[b]	−0.095[b]	−0.171[b]	−0.132[b]	−0.110[c]	−0.105[c]	−0.112[c]	−0.113[c]
	(−2.28)	(−2.10)	(−2.17)	(−2.19)	(−2.13)	(−2.01)	(−2.54)	(−2.00)	(−1.74)	(−1.78)	(−1.87)	(−1.89)
MktRet_1m		−2.482[b]						−2.234[b]				
		(−2.11)						(−2.03)				
MktRet_2m			−2.174[b]						−2.354[b]			
			(−2.26)						(−2.28)			

(Continues)

TABLE 3 (Continued)

	The *upgrade* portfolio						The *downgrade* portfolio					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
MktRet_3m				−2.023[b]						−2.498[b]		
				(−2.02)						(−2.41)		
MktRet_w					−2.165[b]	−2.780[b]					−2.372[b]	−2.812[b]
					(−2.24)	(−2.38)					(−2.20)	(−2.41)
MktVol_1m						2.510[a]						2.468[a]
						(2.78)						(3.05)
YEAR	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
INDUSTRY	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.036	0.243	0.236	0.239	0.240	0.290	0.111	0.217	0.230	0.231	0.211	0.401

Note: This table presents the impact of market conditions, in terms of past market return and volatility, on the performance of the *upgrade* and *downgrade* portfolios. We regress the daily net abnormal returns to the *upgrade* and *downgrade* portfolios, separately, on a set of macroeconomic variables, various market return variables, one market volatility variable, along with two indicator variables of *YEAR* and *INDUSTRY* to control for the potential year and industry fixed effects, respectively. The net abnormal return is calculated as the gross return subtracting the estimated transaction costs for purchasing stocks at 1.5% and for short selling those at 3.0%, multiplied by the corresponding average daily portfolio turnover. The gross return is estimated as the intercept term derived from the CAPM. The four macroeconomic factors are (i) the inflation rate computed from the consumer price index (*CPI*), (ii) the industrial production growth (*PROD*), (iii) the dividend yield (*DIV*), (iv) the detrended short-term interest rate (*IR*). The market return variables are constructed for the one-month (*MktRet_1m*), two-month (*MktRet_2m*), and three-month (*MktRet_3m*) periods before the portfolio is rebalanced as the buy-and-hold return on the FTSE All-Share index, respectively. The three-month weighted market return variable (*MktRet_w*) is constructed as a weighted average buy-and-hold return of the corresponding market index return in the three months before the portfolio is rebalanced. The weights are three for the most recent month, two for the previous month, and one for the third month before the portfolio is rebalanced (Derrien & Womack, 2003). The market volatility variable (*MktVol_1m*) is measured as the standard deviation of the daily return of the relevant FTSE All-Share Index in the month before the portfolio is rebalanced. The performance of the *downgrade* portfolio is measured as the signed net abnormal returns multiplied by −1. The White heteroscedasticity-consistent *t*-statistics are reported in parenthesis. [a], [b], and [c] denote statistical significance at the 1, 5, and 10% levels, respectively.

and

$$\begin{aligned}
[Net\ Abnormal\ Returns]_t &= \alpha_B + \beta_B\left(R_{m,t} - R_{f,t}\right) + \varepsilon_{B,t} \quad \text{if } MktVol_1m_t \geq x; \\
[Net\ Abnormal\ Returns]_t &= \alpha_G + \beta_G\left(R_{m,t} - R_{f,t}\right) + \varepsilon_{G,t} \quad \text{if } MktVol_1m_t < x,
\end{aligned} \tag{8}$$

where x is the threshold value of market return (*MktRet_w*) or market volatility (*MktVol_1m*) to be estimated by the maxim and of the likelihood ratio statistics over all permissible values; the subscripts of $_B$ and $_G$ represent the *bad* and *good* market conditions, respectively.

The Hansen (2000) non-temporal threshold testing procedure as described in Eqs. (7–8) might ignore the possibility of multiple non-temporal thresholds. Gonzalo and Pitarakis (2002) extend the procedure by applying the multiple structural change analysis of Bai and Perron (1998). The multiple non-temporal threshold models are as follows:

$$\begin{aligned}
[Net\ Abnormal\ Returns]_t &= \alpha_1 + \beta_1\left(R_{m,t} - R_{f,t}\right) + \varepsilon_{1,t} && \text{if } MktRet_w_t \leq x_1; \\
&\vdots \\
[Net\ Abnormal\ Returns]_t &= \alpha_j + \beta_j\left(R_{m,t} - R_{f,t}\right) + \varepsilon_{j,t} && \text{if } x_{j-1} < MktRet_w_t \leq x_j; \\
&\vdots \\
[Net\ Abnormal\ Returns]_t &= \alpha_{k+1} + \beta_k\left(R_{m,t} - R_{f,t}\right) + \varepsilon_{k,t} && \text{if } x_k < MktRet_w_t,
\end{aligned} \tag{9}$$

and

$$\begin{aligned}
[Net\ Abnormal\ Returns]_t &= \alpha_1 + \beta_1\left(R_{m,t} - R_{f,t}\right) + \varepsilon_{1,t} && \text{if } MktVol_1m_t \geq x_1; \\
&\vdots \\
[Net\ Abnormal\ Returns]_t &= \alpha_j + \beta_j\left(R_{m,t} - R_{f,t}\right) + \varepsilon_{j,t} && \text{if } x_{j-1} > MktVol_1m_t \geq x_j; \\
&\vdots \\
[Net\ Abnormal\ Returns]_t &= \alpha_{k+1} + \beta_k\left(R_{m,t} - R_{f,t}\right) + \varepsilon_{k,t} && \text{if } x_k > MktVol_1m_t,
\end{aligned} \tag{10}$$

where k represents the number of thresholds. Specifically, the multiple non-temporal threshold testing procedure begins with the null hypothesis of zero threshold against the alternative of one threshold; if this is rejected, it proceeds to two thresholds, and so on. The threshold testing results are based on the p-values computed by the asymptotically correct bootstrap procedure proposed by Hansen (2000).

The left side of Table 4 reports a statistically significant *Threshold* 1 (0.075; t-stat = 2.77), at the 1% level, but an insignificant *Threshold* 2 (0.051; t-stat = 1.06), indicating the presence of one break, but not of two breaks, when market conditions are proxied as the past market return in the case of the performance of the *upgrade* portfolio. Specifically, the coefficient of β_B is significantly negative, –0.123 (t-stat = –2.49), at the 5% level, when the market return is below the identified *Threshold* 1, while the coefficient of β_G (0.025; t-stat = 0.31) is statistically insignificant when market return is above the threshold. In addition, when market conditions are proxied by the past market volatility, the coefficient of β_B is significantly positive, 0.233 (t-stat = 2.64), at the 1% level, when market volatility is above the identified *Threshold* 1, while the coefficient of β_G (0.022; t-stat = 0.12) is statistically insignificant when market volatility is below the threshold.

Overall, our non-temporal analysis results show that the performance of the *upgrade* portfolio increases in bad market conditions (i.e., lower market return and/or higher market volatility). But, once market conditions become better and exceed a certain level, the negative impact of market conditions on the recommendation performance disappears. Our results confirm the importance of taking non-linearity into account when analyzing the relationship between the time-varying recommendation performance and market conditions. The right side of Table 4 shows similar evidence in the case of the performance of the *downgrade* portfolio, confirming Hypothesis 1 that the average net abnormal return to the *upgrade* (*downgrade*) portfolio is significantly *positive* (*negative*) in bad market conditions, but remains statistically insignificant in the rest of the sample period.

TABLE 4 Non-temporal threshold regression analysis results under the CAPM

	The *upgrade* portfolio		The *downgrade* portfolio	
Market conditions in terms of	*MktRet_w*	*MktVol_1m*	*MktRet_w*	*MktVol_1m*
Threshold 1	0.075[a]	0.011[a]	0.076[a]	0.010[a]
	(2.77)	(2.67)	(2.64)	(2.70)
Threshold 2	0.051	0.007	0.046	0.006
	(1.06)	(0.66)	(0.89)	(0.53)
Constant (Bad)	0.006	0.007	0.004	0.005
	(0.57)	(0.45)	(0.36)	(0.63)
Constant (Good)	0.013[c]	0.011b	0.016[b]	0.010[c]
	(1.85)	(2.04)	(2.21)	(1.83)
β_B	–0.123[b]	0.233[a]	–0.085[b]	0.144[b]
	(2.49)	(2.64)	(2.12)	(2.46)
β_G	0.025	0.022	0.029	0.026
	(0.31)	(0.12)	(0.53)	(0.42)
adj. R^2 (Bad)	0.205	0.237	0.172	0.187
adj. R^2 (Good)	0.032	0.029	0.043	0.037

Note: This table presents empirical evidence on whether the dynamic recommendation performance is sensitive to different market conditions, in terms of past market returns or market volatility, using the non-temporal threshold testing procedure, proposed by Hansen (2000) as shown in Eqs. (7–8). The threshold value of market return (*MktRet_w*) or market volatility (*MktVol_1m*) is estimated by the maxim of the likelihood ratio statistics over all permissible values. Specifically, the three-month weighted market return variable (*MktRet_w*) is constructed as a weighted average buy-and-hold return of the corresponding market index return in three months before the portfolio is rebalanced. The weights are three for the most recent month, two for the next, and one for the third month before the portfolio is rebalanced (Derrien & Womack, 2003). The market volatility variable (*MktVol_1m*) is measured as the standard deviation of the daily return of the relevant FTSE All-Share Index in one month before the portfolio is rebalanced. The parameter estimates of β_B and β_G are generated under the CAPM, and the subscripts of $_B$ and $_G$ indicate bad and good market conditions, respectively. The performance of the *downgrade* portfolio is measured as the signed net abnormal returns multiplied by −1. The White heteroskedasticity-consistent *t*-statistics are reported in parenthesis. [a], [b], and [c] denote statistical significance at the 1, 5, and 10% levels, respectively.

5 | ROBUSTNESS CHECKS

In Subsection 5.1, we develop a time-series bootstrap simulation method to distinguish analysts' luck from skill. Our bootstrap simulation results confirm that the observed significant time-specific portfolio performance based on analyst recommendation revisions is not due to random chance, that is, analysts possess sufficient skill to make valuable upgrades and downgrades in bad market conditions. Subsection 5.2 provides very strong evidence showing that our empirical results are robust to various multi-factor assets pricing models, ruling out the concern that our results might be due to a poor model of asset pricing (Barber et al., 2001).

5.1 | Bootstrap simulations

5.1.1 | Time-series bootstrap simulation method

Kosowski, Timmermann, Wermers, and White (2006) develop a *cross-sectional* bootstrap simulation method on mutual funds research, which resamples the residuals from individual fund returns independently, but remains the effect of common risk factors *fixed* historically. Fama and French (2010, p. 1940), however, argue that "failure to account for the joint distribution of fund returns, and of fund and explanatory returns, biases the inferences of Kosowski et al. (2006) toward positive performance". Extending Kosowski et al. (2006), Fama and French (2010) jointly resample both the residuals and risk factors, *ceteris paribus*. Inspired by Kosowski et al. (2006) and Fama and French (2010), we develop a rolling window based bootstrap simulation method to distinguish analysts' luck from skill. Note that our method measures the performance distribution of the best performing rolling windows not only by resampling from the distribution of the *ex-post* best performing rolling windows, but using the information about luck represented by *all* rolling windows. This is a major difference between our method and those employed in previous studies, which generally ignore the possibility that luck distribution encountered by all other performance distributions also provides highly valuable and relevant information (see, e.g., Cuthbertson, Nitzsche, & O'Sullivan, 2008; White, 2000). Our rolling window based bootstrap simulation method thus allows for a comprehensive investigation into the time-varying recommendation performance after explicitly controlling for luck and alleviating the potential bias from misspecification.

First, in the first one-year rolling window of January 3, 1995 to December 29, 1995, we run the CAPM to calculate the estimated alphas, factor loadings, and residuals, using the time series of daily excess returns for the portfolio, $\{(R_{p,t} - R_{f,t});\ t = T_{p,1}, \ldots, T_{p,252}\}$, where $T_{p,1}$ and $T_{p,252}$ are the first and last trading dates, respectively, in the rolling window:

$$R_{p,t} - R_{f,t} = \hat{\alpha}_p + \hat{\beta}_p \left(R_{m,t} - R_{f,t}\right) + \hat{\varepsilon}_{p,t}. \tag{11}$$

Second, we save the coefficient estimates, $\{\hat{\alpha}_p,\ \hat{\beta}_p\}$, the time series of estimated residuals, $\{\hat{\varepsilon}_{p,t};\ t = T_{p,1}, \ldots,\ T_{p,252}\}$, and the t-statistic of alpha, $\hat{t}_{\hat{\alpha}_p}$.

Third, we generate a pseudo-time series of resampled residuals $\{\hat{\varepsilon}^b_{p,t_b};\ t_b = T^b_{p,1}, \ldots,\ T^b_{p,252}\}$ by randomly drawing residuals from the saved residual vector $\{\hat{\varepsilon}_{p,t}\}$ with replacements, where b is the bootstrap simulation index. In the same way, we generate a pseudo-time series of market risk $\{(R_{m,t_b} - R_{f,t_b})^b\}$ by randomly drawing market risk from the original risk factor vector $\{(R_{m,t} - R_{f,t})\}$ with replacements.

Fourth, we generate a time series of pseudo-daily excess returns $(R_{p,t} - R_{f,t})^b$ in the rolling window, imposing the null hypothesis of zero true performance ($\alpha_p = 0$):

$$\left\{ \left(R_{p,t} - R_{f,t}\right)^b = 0 + \hat{\beta}_p \left(R_{m,t_b} - R_{f,t_b}\right)^b + \hat{\varepsilon}^b_{p,t_b} \right\}, \tag{12}$$

where $t = T_{p,1}, \ldots, T_{p,252};\ t_b = T^b_{p,1}, \ldots,\ T^b_{p,252}$.

Finally, we regress the pseudo-daily excess returns $(R_{p,t} - R_{f,t})^b$ on the market factor:

$$\left(R_{p,t} - R_{f,t}\right)^b = \hat{\alpha}^b_p + \hat{\beta}_p \left(R_{m,t} - R_{f,t}\right) + \hat{\varepsilon}_{p,t}. \tag{13}$$

The simulated $\hat{\alpha}^b_p$ represents the sampling variation around zero true performance, entirely due to random chance (luck). Repeating the above steps in each of the 4,420 rolling windows over the period January 1996 to June 2013, we obtain a time series of simulated alphas, $\{\hat{\alpha}^b_p\}$, and their corresponding t-statistics, $\{\hat{t}_{\hat{\alpha}^b_p}\}$. We then order all simulated $\hat{\alpha}^b_p$ into a cumulative distribution function (CDF) of simulated $\hat{\alpha}^b_p$–a separate time series of *luck* distribution from the worst performing rolling window to the best performing rolling window, all of which are completely due to analysts' luck rather than skill. In this study, we repeat the above bootstrap simulation 10,000 times, say, $b = 1, \ldots, 10{,}000$.

5.1.2 | Bootstrap simulation results

We focus on presenting the distribution of t-statistics of the net abnormal returns to the *upgrade* and *downgrade* portfolios, as the t-statistic scales the net abnormal return by its standard errors and thus has superior statistical properties (see, Fama & French, 2010; Meyer et al., 2012).[15] Specifically, we compare the values of the t-statistics at selected percentiles of CDFs of the actual t-statistics with the averages of the 10,000 simulated t-statistics at the same percentiles. We reject the null hypothesis that the portfolio performance is due to random chance (at the 95% confidence level) and infer that analysts have skill at this percentile when the actual t-statistic is higher than the simulated t-statistic in 95%

TABLE 5 The percentiles of *t*-statistics of the actual and simulated net abnormal returns to the *upgrade* and *downgrade* portfolios under the CAPM

	The *upgrade* portfolio			The *downgrade* portfolio		
%	Simulated *t*-stat	Actual *t*-stat	% (Simulated < Actual)	Simulated *t*-stat	Actual *t*-stat	% (Simulated < Actual)
1	−0.52	−0.82	1.87	−0.78	−0.95	1.92
2	−0.11	−0.47	2.95	−0.35	−0.58	2.95
3	0.15	−0.26	3.44	−0.09	−0.35	3.63
4	0.29	−0.11	4.54	0.08	−0.21	4.54
5	0.38	−0.01	5.77	0.20	−0.11	6.19
10	0.64	0.31	6.56	0.52	0.16	7.37
20	0.87	0.65	8.50	0.78	0.51	9.42
30	1.02	0.85	14.68	0.92	0.70	14.61
40	1.14	1.01	17.63	1.01	0.81	20.48
50	1.26	1.18	24.17	1.06	0.92	31.63
60	1.37	1.35	37.47	1.10	1.02	50.74
70	1.50	1.54	55.82	1.15	1.16	54.66
80	1.66	1.77	61.66	1.26	1.37	68.87
90	1.94	2.10	68.07	1.48	1.72	81.03
95	2.25	2.41	77.24	1.75	2.06	86.30
96	2.35	2.51	83.71	1.85	2.17	88.16
97	2.51	2.68	89.23	2.00	2.36	91.90
98	2.73	2.89	92.70	2.18	2.57	93.97
99	2.94	3.16	95.71	2.57	3.11	96.47

Note: This table presents the values of *t*-statistics at selected percentiles (%) of the distribution of *t*-statistics of the actual and simulated net abnormal returns to the *upgrade* and *downgrade* portfolios, as well as the percentage of the 10,000 simulation runs that produce lower values of *t*-statistics at the selected percentiles than those actual net abnormal returns (% Simulated < Actual) over the period January 1996 to June 2013. The net abnormal return is calculated as the gross return subtracting the estimated round-trip transaction costs for purchasing stocks at 1.5% and for short selling those at 3.0%, multiplied by the corresponding average daily portfolio turnover. The gross return is estimated as the intercept term derived from the CAPM. The performance of the *downgrade* portfolio is measured as the signed net abnormal returns multiplied by −1.

or more of the 10,000 simulations. Our simulation results in Table 5 mirror the net abnormal returns to the *upgrade* and *downgrade* portfolios under the CAPM.

The left side of Table 5 shows that the left tail percentiles of the actual *t*-statistics of the net abnormal returns to the *upgrade* portfolio are below the corresponding average values from the simulations. For example, the 5th and 10th percentiles of the actual *t*-statistics, −0.01 and 0.31, respectively, are much lower than the corresponding average simulated values of 0.38 and 0.64, suggesting that the relatively poor portfolio performance in the left tails over some time periods is due to analysts' inferior skill rather than their poor luck. Thus, the left tails of the CDFs of *t*-statistics indicate the existence of some time periods, during which analysts' upward revisions could not result in positive true alpha relative to the passive benchmark. Thus, it is unlikely that investors could make profits by purchasing stocks with upward revisions in these time periods.

However, the right tails of the CDFs of *t*-statistics suggest that, on average, analysts have skill in making valuable upward revisions that can cover the size of transaction costs over some time periods (e.g., in bad market conditions). Specifically, the right tail percentiles of the actual *t*-statistics are always higher than the corresponding average simulated values for all percentiles above the 70th. For example, the 90th (95th) percentile of the actual *t*-statistics is 2.10

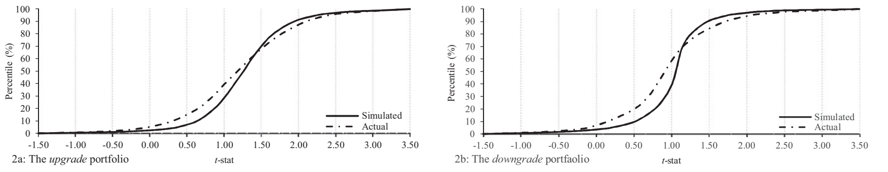

FIGURE 2 The actual and simulated cumulative density functions (CDFs) of t-statistics of the net abnormal returns to the *up/downgrade* portfolio, under the CAPM

(2.41), much higher than the corresponding average simulated value of 1.94 (2.25). In particular, under the CAPM, the 99th percentile of the actual t-statistics is higher than the corresponding simulated t-statistics at the 95% confidence level, confirming the existence of some time periods during which upward revisions made by analysts can generate statistically significant net abnormal returns. We further confirm that a vast majority of these time periods showing the significant recommendation performance happen in bad market conditions. Similarly, Figure 2a illustrates the CDFs of the actual t-statistics of the net abnormal returns to the *upgrade* portfolio and the corresponding CDFs from the simulations, under the CAPM.

The right side of Table 5 shows that the left tail percentiles of the actual t-statistics of the net abnormal returns to the *downgrade* portfolio are still to the left of the corresponding average values from the simulations, irrespective of the asset pricing models employed. For example, the 5th and 10th percentiles of the actual t-statistics are −0.11 and 0.16, respectively, much lower than the corresponding average simulated values of 0.20 and 0.52, again suggesting that the relatively poor portfolio performance in the left tails is due to analysts' inferior skill over some time periods rather than their poor luck. Therefore, the left tails of the CDFs of t-statistics suggest that it is unlikely that investors can make profits by short selling stocks with downgrade revisions over these time periods.

In contrast, the right tail percentiles of the actual t-statistics suggest that, on average, analysts have skill in making valuable downward revisions over some time periods. Specifically, the CDF of the actual t-statistics under the CAPM moves to the right of the average values at about the 70th percentile from the simulations. For example, the 90th and 95th percentiles of the actual t-statistics are 1.72 and 2.06, respectively, higher than the corresponding average simulated values of 1.48 and 1.75. In particular, the 99th percentile of the actual t-statistics are higher than the corresponding simulated t-statistics at the 95% confidence level, confirming the existence of some time periods, during which analysts show skill in making valuable downward revisions; almost all of these time periods showing the significant recommendation performance happen in bad market conditions. Similarly, Figure 2b illustrates the CDFs of the actual t-statistics of the net abnormal returns to the *downgrade* portfolio and the corresponding CDFs from the simulations, under the CAPM.

In summary, our bootstrap simulations suggest that financial analysts have sufficient skills in making valuable upward and downward revisions over certain time periods, i.e., bad market conditions, rather than during the whole sample period.

5.2 | Alternative multi-factor asset pricing models

We also estimate the gross returns to the *up/downgrade* portfolio using the Fama and French (1993) three-factor model (hereafter, the FF model) and the Carhart (1997) four-factor model (hereafter, the FFC model):

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_p \left(R_{m,t} - R_{f,t}\right) + s_p SMB_t + h_p HML_t + \varepsilon_{p,t}; \tag{14}$$

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_p \left(R_{m,t} - R_{f,t}\right) + s_p SMB_t + h_p HML_t + m_p MOM_t + \varepsilon_{p,t}, \tag{15}$$

where SMB_t, HML_t, and MOM_t represent the daily returns on zero-investment factor-mimicking portfolios for size, book-to-market (B/M), and price momentum, respectively.[16]

Using the FF and FFC models, we further replicate all empirical investigations in Section 4 and bootstrap simulations in Subsection 5.1, confirming that our results are qualitatively unchanged. These robustness results are provided as supplementary materials (see Tables A1–A3 and Figures A1–A4).

6 | CONCLUSIONS

This study examines the time-varying performance of analyst recommendation revisions in the UK, using a comprehensive sample of 70,220 analyst recommendation revisions over the period January 1995 to June 2013. We construct an *upgrade* portfolio, consisting of all stocks with upward revisions to Strong Buy or Buy recommendations from previous Strong Sell, Sell, or Hold recommendations, as well as a *downgrade* portfolio, consisting of all stocks with downward revisions to Strong Sell, Sell, or Hold recommendations from previous Strong Buy or Buy recommendations. Specifically, the *upgrade* and *downgrade* portfolios are updated daily; for each upward or downward revision, the recommended stock enters the *upgrade* or *downgrade* portfolio at the close of trading on the day the revision is announced, and then remains in the portfolio for up to five trading days.

First, we find that the performance of the *upgrade* and *downgrade* portfolios varies considerably over time on a rolling-window basis. In particular, the *upgrade* (*downgrade*) portfolio generates significantly positive (negative) net abnormal returns in bad market conditions, e.g., the *dot-com* bubble burst in 2000 and the *credit* crisis in 2007. In addition, we find that the time-varying portfolio performance is driven by market conditions as measured in terms of both past market return and market volatility. Specifically, the *up/downgrade* portfolio leads to superior performance in bad conditions, i.e., in periods where there are lower market return and/or higher market volatility. Subsequently, we employ the Hansen (2000) non-temporal threshold regression model to test whether the dynamic recommendation performance is sensitive to different market conditions. Our non-temporal analysis results confirm that significantly positive (negative) net abnormal returns to the *upgrade* (*downgrade*) portfolio in bad market conditions, while once market conditions become better and exceed a certain level, the observed impact of market conditions on the recommendation performance disappears.

Our results hold up against an array of robustness checks, including bootstrap simulations and various multi-factor asset pricing models. Our bootstrap simulations confirm that the observed significant time-specific recommendation performance is due to analysts' skill rather than luck (i.e., random chance). That is, analysts possess sufficient skill to make valuable upgrades and downgrades over certain time periods, which are related to changing market conditions. Also, our results are robust to various single- and multi-factor assets pricing models, ruling out Barber et al. (2001) concern that the observed significant recommendation performance could be due to a poor model of asset pricing.

In summary, we find strong evidence from the UK market broadly supporting Loh and Stulz (2018) that analyst recommendations are more valuable in bad times than good. Future research might seek to further confirm the generality of findings in other market settings and also to investigate the detailed reasons for the effect.

ACKNOWLEDGEMENTS

The authors are grateful for comments and suggestions from the Editor (Anthony Sounders), Christopher J. Adcock, Joao F. Gomes, Iftekhar Hasan, Brian M. Lucey, Haluk Unal, and Rossen I. Valkanov. Su also thanks financial supports from Newcastle University Business School under the School Strategic Fund 2017/18 and Faculty of Humanities and Social Sciences, Newcastle University (HaSS) under the Faculty Research Fund – Autumn 2019 (OSR/0372/FR19/0001). An early version of this paper was circulated under the title "Are Markets Semi-Strong-Form Efficient or Adaptive? Evidence on the value of Analyst Recommendations in the UK". All remaining errors and omissions are our own.

ENDNOTES

[1] For example, the ring fencing requirements of the Vickers Report lead to a significant restructuring of the UK banking system (Goodhart & Schoenmaker, 2016). See more discussions on the differences of market structure in the two markets in the speech of *'Comparing UK and US Macroprudential Systems: Lessons for China'* given by Donald Kohn at the Global Financial Forum, Tsinghua University, Beijing, on 11 May, 2014 (available at: https://goo.gl/SHWkdX).

[2] Extant analyst research generally conducts event-time analysis, but these event studies fail to capture the possibility that markets evolve over time. In addition, Barber et al. (2001) point out that event-time analysis does not measure the profits to an implementable investment strategy (see McWilliams, Siegel, & Teoh, 1999 for more discussions on the problems and associated solutions related to event studies).

[3] In addition, Jegadeesh and Kim (2006) report that stock prices react significantly to analyst recommendation revisions in the Group of Seven (G7) industrialized counties from November 1993 to July 2002, except for Italy. They further point out that US analysts are more skilled at identifying mispriced stocks than their counterparts in the G7 counties, probably due to the better compensation structure in the US market.

[4] In addition, Jegadeesh and Kim (2006) report that stock prices react significantly to analyst recommendation revisions in the Group of Seven (G7) industrialized counties from November 1993 to July 2002, except for Italy. They further point out that US analysts are more skilled at identifying mispriced stocks than their counterparts in the G7 counties, probably due to the better compensation structure in the US market.

[5] Reactions to analyst recommendation revisions during the first few days after their announcements (see, e.g., Stickel, 1995; Womack, 1996; Green, 2006).

[6] If a stock is recommended by more than one brokerage house on a given date, then that stock will appear multiple times in the portfolio on that date, once for each brokerage house (see, Barber, Lehavy, & Trueman, 2007). Specifically, in our sample, a very small proportion (4.29%) of analyst recommendation revisions is made by more than one BH on a given date, e.g., 3.24% (832 out of 25,701) of upgrades and 5.19% (1,575 out of 30,374) downgrades.

[7] Like Mokoaleli-Mokoteli, Taffler, and Agarwal (2009), we exclude 4,297 upward revisions and 4,847 downward revisions on utilities and financials (the two-digit ICB codes 75, 83, and 85–87) due to their highly regulated nature, and replicate all empirical analyses in Sections 4 and 5, showing qualitatively similar results, which are not reported to save space, but available on request. In fact, it is a common practice for the analyst research not to exclude financials and/or utilities (see, e.g., Barber et al., 2001, 2007; Green, 2006, Fang & Yasuda, 2014), as financial analysts often make stock recommendations focusing on one particular industry or sector.

[8] We explicitly exclude the return on the first trading day, as many investors, particularly small investors, tend to react to information with a delay. Barber et al. (2001, p. 534) argue that "it is impractical for them to engage in the daily portfolio rebalancing that is needed to respond to the changes".

[9] The value-weighted returns enable us to better capture the economic significance of our results, while the equal-weighted returns are, on average, biased upward due to the bid-ask bounce, that is, the returns of large size firms will be more heavily represented in the aggregate returns than those of small size firms (see, Barber et al., 2001).

[10] Similarly, Barber et al. (2001) estimate the average round-trip transaction costs of 1.31% in the US. Despite the lack of readily available data regarding short selling costs in the UK, we assume a short selling cost of 3.0%, according to Su, Zhang, Bangassa, and Joseph (2019).

[11] In fact, if markets are truly efficient, the t-statistics of the net abnormal returns should be within the 95% confidence interval (p-value ≥ 0.05) over the entire sample period, that is, the absolute values of t-statistics should be consistently less than the critical value of 1.96 (Kim et al., 2011).

[12] Our results are qualitatively unchanged when the recommended stocks remain in the *up/downgrade* portfolio for three trading days, suggesting that early access to analyst recommendation revisions could generate incremental investment value (see, also, Li, 2005; Green, 2006). Not surprisingly, the reported significant net abnormal returns to the *up/downgrade* portfolio disappear when the recommended stocks remain in the portfolio until the recommendations are changed or expired (an average of 68 days as shown in Appendix A), consistent with prior analyst research. For example, in Barber et al. (2007), the recommended stock enters the appropriate portfolio at the close of trading on the day the stock recommendation is released, and remains in the portfolio until the recommendation is either up/downgraded or dropped from coverage by the brokerage house, while such investment strategies generate no consistently profitable performance after taking transaction costs into account (see, also, Barber et al., 2001; Jegadeesh, Kim, Krische, & Lee, 2004; Mikhail, Walther, & Willis, 2004; Hall & Tacon, 2010).

[13] In Tables 3–5, we present the performance of the *downgrade* portfolio as the signed net abnormal returns multiplied by −1 to make direct comparison with the performance of the *upgrade* portfolio.

[14] Specifically, we first estimate the threshold parameter and the coefficient estimates to obtain the associated residuals; we then re-order the residuals temporally and perform the test for serial correlations using the re-ordered residuals. Therefore, the inclusion of only the first lag of the dependent variable is sufficient to ensure that the residuals do not exhibit serial correlation. The threshold test results are based on the *p*-values measured by the asymptotically correct bootstrap procedure proposed by Hansen (2000).

[15] Specifically, we use the Newey–West heteroskedasticity-robust standard errors to calculate *t*-statistics. Like Meyer et al. (2012), we do not account for autocorrelation for two main reasons. First, the majority of rolling window regressions do not report autocorrelation at the 5% significance level using the Breusch-Godfrey test. Second, it has the advantage of enhancing the comparability between simulated and actual *t*-statistics through a uniform test specification. This is because our time-series bootstrap simulations consist of random drawings of individual daily returns with replacements, which means the time series drawn cannot contain any true underlying autocorrelation by design.

[16] The daily returns on size, value, and momentum in the UK stock market are collected from the Xfi Centre for Finance and Investment at University of Exeter.

REFERENCES

Amiram, D., Landsman, W. R., Owens, E. L., & Stubben, S. (2018). How are analysts' forecasts affected by high uncertainty? *Journal of Business Finance & Accounting*, *45*, 295–318.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, *66*, 47–78.

Barber, B. M., Lehavy, R., McNichols, M., & Trueman, B. (2001). Can investors profit from the prophets? Security analyst recommendations and stock returns. *Journal of Finance*, *56*, 531–563.

Barber, B. M., Lehavy, R., McNichols, M., & Trueman, B. (2003). Reassessing the returns to analysts' stock recommendations. *Financial Analyst Journal*, *59*, 88–96.

Barber, B. M., Lehavy, R., & Trueman, B. (2007). Comparing the stock recommendation performance of investment banks and independent research firms. *Journal of Financial Economics*, *85*, 490–517.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, *77*, 623–685.

Boni, L., & Womack, K. L. (2006). Analysts, industries, and price momentum. *Journal of Financial and Quantitative Analysis*, *41*, 85–109.

Bradley, D., Gokkaya, S., & Liu, X. (2017). Before an analyst becomes an analyst: Does industry experience matter? *Journal of Finance*, *72*, 751–792.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, *52*, 57–82.

Chan, K.-S., Hansen, B. E., & Timmermann, A. (2017). Guest editors' introduction: Regime switching and threshold models. *Journal of Business and Economic Statistics*, *35*, 159–161.

Chang, J. W., & Choi, H. M. (2017). Analyst optimism and incentives under market uncertainty. *Financial Review*, *52*, 307–345.

Chopra, V. K. (1998). Why so much error in analysts' earnings forecasts? *Financial Analysts Journal*, *54*, 35–42.

Cowles, A. (1933). Can stock market forecasters forecast? *Econometrica*, *1*, 309–324.

Cuthbertson, K., Nitzsche, D., & O'Sullivan, N. (2008). UK mutual fund performance: Skill or luck? *Journal of Empirical Finance*, *15*, 613–634.

Derrien, F., & Womack, K. L. (2003). Auctions vs. bookbuilding and control of under-pricing in hot IPO markets. *Review of Financial Studies*, *16*, 31–61.

Dimson, E., & Fraletti, P. (1986). Brokers' recommendations: The value of a telephone tip. *Economic Journal*, *96*, 139–159.

Dimson, E., & Marsh, P. (1984). An analysis of brokers' and analysts' unpublished forecasts of UK stock returns. *Journal of Finance*, *39*, 1257–1292.

Emery, D. R., & Li, X. (2009). Are the Wall Street analyst rankings popularity contests? *Journal of Financial and Quantitative Analysis*, *44*, 411–438.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, *49*, 283–306.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*, 3–56.

Fama, E. F., & French, K. R. (2010). Luck versus skill in the cross-section of mutual fund returns. *Journal of Finance*, *65*, 1915–1947.

Fang, L., & Yasuda, A. (2009). The effectiveness of reputation as a disciplinary mechanism in sell-side research. *Review of Financial Studies*, *22*, 3735–3777.

Fang, L., & Yasuda, A. (2014). Are stars' opinions worth more? The relation between analyst reputation and recommendation values. *Journal of Financial Services Research*, *46*, 235–269.

Flannery, M. J., & Protopapadakis, A. A. (2002). Macroeconomic factors do influence aggregate stock returns. *Review of Financial Studies*, *15*, 751–782.

Forbes, W. P., Murphy, A., O'Keeffe, C., & Su, C. (2020), Are financial analysts eager postmen of bubble psychology? Evidence in the United Kingdom. *International Journal of Finance and Economics*, *24*, 120–137.

Gonzalo, J., & Pitarakis, J.-Y. (2002). Estimation and model selection based inference in single and multiple threshold models. *Journal of Econometrics*, *110*, 319–352.

Goodhart, C. A. E., & Schoenmaker, D. (2016). The global investment banks are now all becoming American: Does that matter for Europeans? *Journal of Financial Regulation*, *2*, 163–181.

Green, T. C. (2006). The value of client access to analyst recommendations. *Journal of Financial and Quantitative Analysis*, *41*, 1–24.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, *70*, 393–408.

Hall, J. L., & Tacon, P. B. (2010). Forecast accuracy and stock recommendations. *Journal of Contemporary Accounting and Economics*, *6*, 18–33.

Hansen, B. E. (2000). Sample splitting and threshold estimation. *Econometrica*, *68*, 575–603.

Hjalmarsson, E. (2010). Predicting global stock returns. *Journal of Financial and Quantitative Analysis*, *45*, 49–80.

Hudson, R. S., Dempsey, M., & Keasey, K. (1996). A note on the weak form efficiency of capital markets: The application of simple technical trading rules to UK stock prices – 1935 to 1994. *Journal of Banking and Finance*, *20*, 1121–1132.

Irvine, P., Lipson, M., & Puckett, A. (2007). Tipping. *Review of Financial Studies*, *20*, 741–768.

Ivković, Z., & Jegadeesh, N. (2004). The timing and value of forecast and recommendation revisions. *Journal of Financial Economics*, *73*, 433–463.

Jegadeesh, N., Kim, J., Krische, S. D., & Lee, C. M. C. (2004). Analyzing the analysts: When do recommendations add value? *Journal of Finance*, *59*, 1083–1124.

Jegadeesh, N., & Kim, W. (2006). Value of analyst recommendations: International evidence. *Journal of Financial Markets*, *9*, 274–309.

Jegadeesh, N., & Kim, W. (2010). Do analysts herd? An analysis of recommendations and market reactions. *Review of Financial Studies*, *23*, 901–937.

Kacperczyk, M., & Seru, A. (2007). Fund manager use of public information: New evidence on managerial skills. *Journal of Finance*, *62*, 485–528.

Kacperczyk, M., Van Nieuwerburgh, S., & Veldkamp, L. (2016). A rational theory of mutual funds' attention allocation. *Econometrica*, *84*, 571–626.

Keim, D. B., & Madhavan, A. (1998). The cost of institutional equity trades. *Financial Analysts Journal*, *54*, 50–69.

Kosowski, R., Timmermann, A., Wermers, R., & White, H. A. L. (2006). Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis. *Journal of Finance*, *61*, 2551–2595.

Kucheev, Y. O., Ruiz, F., & Sorensson, T. (2017). Do stars shine? Comparing the performance persistence of star sell-side analysts listed by Institutional Investor, the Wall Street Journal, and StarMine. *Journal of Financial Services Research*, *52*, 277–305.

Li, X. (2005). The persistence of relative performance in stock recommendations of sell-side financial analysts. *Journal of Accounting and Economics*, *40*, 129–152.

Loh, R. K., & Stulz, R. M. (2011). When are analyst recommendation changes influential? *Review of Financial Studies*, *24*, 593–627.

Loh, R. K., & Stulz, R. M. (2018). Is sell-side research more valuable in bad times? *Journal of Finance*, *73*, 959–1013.

McWilliams, A., Siegel, D., & Teoh, S. H. (1999). Issues in the use of the event study methodology: A critical analysis of corporate social responsibility. *Organizational Research Methods*, *2*, 340–365.

Mehran, H., & Stulz, R. M. (2007). The economics of conflicts of interest financial institutions. *Journal of Financial Economics*, *85*, 267–598.

Merkley, K., Michaely, R., & Pacelli, J. (2017). Does the scope of the sell-side analyst industry matter? An examination of bias, accuracy, and information content of analyst reports. *Journal of Finance*, *72*, 1285–1334.

Meyer, S., Schmoltzi, D., Stammschulte, C., Kaesler, S., Loos, B., & Hackethal, A. (2012). Just unlucky? – A bootstrapping simulation to measure skill in individual investors' investment performance. Unpublished paper, Goethe University Frankfurt.

Mikhail, M. B., Walther, B. R., & Willis, R. H. (2004). Do security analysts exhibit persistent differences in stock picking ability? *Journal of Financial Economics*, *74*, 67–91.

Mokoaleli-Mokoteli, T., Taffler, R. J., & Agarwal, V. (2009). Behavioural bias and conflicts of interest in analyst stock recommendations. *Journal of Business Finance and Accounting*, *36*, 384–418.

Ryan, P., & Taffler, R. J. (2006). Do brokerage houses add value? The market impact of UK sell-side analyst recommendation changes. *British Accounting Review*, *38*, 371–386.

Shen, C.-H., & Chih, H.-L. (2009). Conflicts of interest in the stock recommendations of investment banks and their determinants. *Journal of Financial and Quantitative Analysis*, *44*, 1149–1171.

Stickel, S. E. (1995). The anatomy of the performance of buy and sell recommendations. *Financial Analysts Journal*, *51*, 25–39.

Su, C., Zhang, X., Bangassa, K., & Joseph, N. L. (2019). On the investment value of sell-side analyst recommendation revisions in the UK. *Review of Quantitative Finance and Accounting*, *53*, 257–293.
Timmermann, A., & Granger, C. W. J. (2004). Efficient market hypothesis and forecasting. *International Journal of Forecasting*, *20*, 15–27.
Trueman, B. (1994). Analyst forecasts and herding behaviour. *Review of Financial Studies*, *7*, 97–124.
White, H. (2000). A reality check for data snooping. *Econometrica*, *68*, 1097–1126.
Womack, K. L. (1996). Do brokerage analysts' recommendations have investment value? *Journal of Finance*, *51*, 137–167.

AUTHOR BIOGRAPHIES

Chen Su has held the posts of Lecturer and Senior Lecturer in Corporate Finance since joining Newcastle University Business School in September 2011. Prior to this he earned a bachelor's degree in Mechanical Engineering (Hons) from Tongji University, Shanghai, followed by an MSc in Finance and a PhD in Finance from the University of Liverpool.

Hanxiong Zhang is a Lecturer in Financial Management at Sheffield University Management School. He holds a PhD in Financial Economics from Newcastle University. Before joining Sheffield University Management School in March 2019, he worked as a Senior Lecturer in Banking and Finance at Lincoln International Business School, the University of Lincoln. Hanxiong is a CFA charterholder.

Robert S. Hudson is a Professor of Finance at Hull University Business School. Prior to this post, Robert was a Professor of Finance and Head of Economics Group at Newcastle University Business School and, prior to that, a Senior Lecturer at Leeds University Business School. He has extensive business and consultancy experience and before entering the academic world, worked in the financial services industry for a number of years.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Su C, Zhang H, Hudson RS. The time-varying performance of UK analyst recommendation revisions: Do market conditions matter? *Financial Markets, Inst. & Inst.* 2020;29:65–89. https://doi.org/10.1111/fmii.12126

APPENDIX A: THE DISTRIBUTION OF UK ANALYST RECOMMENDATIONS

						Analyst recommendation frequency									
						Strong Buys (1 & 2)		Buys (3 & 4)		Holds (5)		Sells (6 & 7)		Strong Sells (8 & 9)	
Year	No. of firms covered	No. of brokerage houses	Average duration	Average rating	No. of analyst recommendations	No.	%	No.	%	No.	%	No.	%	No.	%
1995	861	28	92.7	2.49	12,516	4,508	36.02	702	5.61	5,526	44.15	263	2.10	1,517	12.12
1996	1,041	41	83.0	2.35	21,289	8,234	38.68	1,687	7.92	9,000	42.28	502	2.36	1,866	8.77
1997	1,172	50	68.9	2.21	29,988	13,179	43.95	2,520	8.40	11,409	38.05	676	2.25	2,204	7.35
1998	1,220	52	60.3	2.23	29,939	12,487	41.71	3,826	12.48	10,214	34.12	1,213	4.05	2,199	7.34
1999	1,167	50	67.9	2.14	26,274	11,308	43.04	4,046	15.40	8,373	31.87	997	3.79	1,550	5.90
2000	1,102	56	69.2	2.01	22,090	10,325	46.74	3,647	16.51	6,560	29.70	633	2.87	925	4.19
2001	1,100	52	73.8	2.33	19,819	7,639	38.54	2,545	12.84	6,855	34.59	1,056	5.33	1,724	8.70
2002	1,061	49	83.7	2.31	17,928	7,467	41.65	2,041	11.38	5,560	31.01	1,146	6.39	1,714	9.56
2003	1,019	55	66.0	2.26	20,250	8,768	43.30	2,225	10.99	6,264	30.93	1,125	5.56	1,868	9.22
2004	1,038	60	62.0	2.26	23,621	10,096	42.74	2,725	11.54	7,546	31.95	1,096	4.64	2,158	9.14
2005	1,111	59	61.2	2.37	24,976	9,743	39.01	2,737	10.96	8,572	34.32	1,302	5.21	2,622	10.50
2006	1,174	60	61.1	2.20	23,775	10,781	45.35	2,514	10.57	7,275	30.60	1,237	5.20	1,968	8.28
2007	1,169	55	63.8	2.10	18,629	9,151	49.12	1,972	10.59	5,398	28.98	742	3.98	1,366	7.33

Year	No. of firms covered	No. of brokerage houses	Average duration	Average rating	No. of analyst recommendations	Analyst recommendation frequency									
						Strong Buys (1 & 2)		Buys (3 & 4)		Holds (5)		Sells (6 & 7)		Strong Sells (8 & 9)	
						No.	%	No.	%	No.	%	No.	%	No.	%
2008	1,153	62	63.5	2.16	19,370	9,776	50.47	1,536	7.93	5,329	27.51	696	3.59	2,033	10.50
2009	1,034	60	61.7	2.18	21,417	10,318	48.18	1,987	9.28	6,208	28.99	685	3.20	2,219	10.36
2010	1,003	44	70.3	1.92	16,794	9,616	57.26	1,366	8.13	4,440	26.44	265	1.58	1,107	6.59
2011	964	44	74.0	1.95	16,389	9,175	55.98	1,398	8.53	4,440	27.09	254	1.55	1,122	6.85
2012	947	44	69.2	1.97	13,587	7,521	55.35	1,066	7.85	3,842	28.28	169	1.24	989	7.28
2013 (January to June)	724	33	69.8	2.12	5,514	2,741	49.71	413	7.49	1,791	32.48	69	1.25	500	9.07
Overall	**2,905**	**144**	**68.0**	**2.20**	**384,165**	**172,833**	**44.99**	**40,953**	**10.66**	**124,602**	**32.43**	**14,126**	**3.68**	**31,651**	**8.24**

Note: This appendix presents the distribution of 384,165 UK analyst recommendations in each year over the period January 1995 to June 2013, in terms of the number of recommended firms, the number of brokerage houses, as well as the average rating and number of analyst recommendations. All real-time analyst recommendations are obtained from *Morningstar Company Intelligence*. A rating of 1 reflects a strong buy, 2 a buy, 3 a weak buy, 4 a weak buy/hold, 5 a hold, 6 a hold/sell, 7 a weak sell, 8 a sell, and 9 a strong sell, which are reclassified into five categories: Strong Buy (1 & 2), Buy (3 & 4), Hold (5), Sell (6 & 7), and Strong Sell (8 & 9). We report the average rating for analyst recommendations in each year based on the five-point rating scale. The average duration reflects the gap between the starting and expiration dates of analyst recommendations.